

日本語習得適性テストの問題点と実用性

—今後の研究開発へ向けての序論として—

柿崎 里奈

要 旨

本稿は今後の適性テスト研究の序論であり、適性テスト研究開発への課題を明らかにするために論じたものである。

十年程前まで名古屋大学を中心に開発研究されていた「日本語習得適性テスト」は、日本語の習得能力を視覚、聴覚、文法解析の三つの分野から測定するものだったが、それぞれに問題点が残されていて、まだまだ未完成と言わざるを得ない。これらの問題点を踏まえて、これからの課題や、コンピュータ化も含めた実用性について適性テストの様々な可能性を考えた。

【キーワード】 日本語習得適性テスト, 視覚情報処理能力, 聴覚能力処理能力 文法情報解析能力, 実用性, コンピュータソフトウェア化

1. はじめに

私は今まで数校の就学生対象日本語学校に勤めてきて、学習者達のクラス編成について改善しなければならない部分が多いと感じてきた。どの学校でも、クラスが固定するまでに大体2週間から1ヶ月はかかってしまうのが現状だ。その間、クラス移動を余儀なくされた学習者達の中にはあちこちのクラスをたらい回しにされ、クラス変更のなかった学習者と比べて精神的にダメージを受けてしまう人もいる。また、教師側としても教室に行くたびに若干ではあっても学習者の数が変わっているというのは、準備してきた授業を効率よく進める妨げにもなりかねないということで、学期が始まってからのクラス移動はあまり歓迎されない。

もう1つの問題として、プレースメント・テストによって既習段階だけみてあとはランダムに分けているため、同じクラス内でも学期末には習得能力の高い学習者とそうではない学習者との間に大きく差が開いていることが挙げられる。学期の途中から教師側もその差に気がついてはいても、どちらに照準を合わせた授業を行うかによって、より能力の高い学習者に合わせた場合は能力の低い学習者は目標レベルに達することなく学期を終了し次の学期にも再履修になってしまうということがあるし、能力の低い学習者に合わせた授業を行った場合は能力の高い学習者の能力を十分に引き出せないまま、学期を終了してしまうこともある。

このようなことになってしまう原因の一つは、プレースメント・テストの結果だけでクラス変えを行ってしまうことにあるのではないかと考えた。もし、プレースメント・テスト以外でこのような問題を事前に回避できるテストがあって、各人の習得能力別にクラス分けができれば、学習者にとっても教師にとっても、より効率的な授業ができるようになるのではないかと考え、「日本語習得適性テスト」に注目してみた。

本稿では、名古屋大学を中心に行われた「日本語習得適性テスト」の分析結果から見られた問題点をあげた上で、今後の課題や、実用性が本当にあるのかなどについて考察していく。

2. 先行研究

外国語習得の適性テストの開発は主にアメリカで行われていた。まずは第一次大戦後に書き言葉を対象としたテストが開発され、その後、第二次大戦後 1950 年頃から音声も対象とした適性テストの研究が始まった。代表的なところではキャロル (Carroll, J.B.) の MLAT (Modern Language Aptitude Test) とピムズラー (Pimsleur, P) の PLAB (Pimsleur Language Aptitude Battery) が挙げられる。

キャロルの考える適性とは学習動機とは切り離して考えるもので、外国語を容易に早く学ぶことの出来る資質だとし、一般的な知能とは分けて定義している。1959年に米商務省で外国語訓練対象者を選抜するために、より早く目標に達する能力を測定する目的でサポン (Sapon, S.M.) と共同で開発されたが、今ではその縮小版や児童版等も開発され、多くの場所で活用されている。MLATは以下の4つの要素によって作られている。

- 1) 音声をコード化する能力
- 2) 文法に関する感受性
- 3) 外国語の語句などを機械的に学習する能力
- 4) 帰納的に言語を学習する能力

一方、ピムズラーの PLAB は、MLAT とは少し違った観点から作られている。学習動機が言語適性の重要な部分となり、外国語の適性は一般的な知能と関係しているという考えに基づき、1966 年に開発された。PLAB は以下の要素から成る。

- 1) 言葉の知識や言語使用における分析能力（言語的知能）
- 2) 学習への興味、モチベーション（動機付け）
- 3) 音を聞き分け、コード化する能力（聴解能力）

以上の 3 つの要素に加えて、のみこみの速さというものを総合的な知的能力から測定するために、外国語以外の主要教科の成績も指標に含んでいるところが特徴的なテストである。

これらを踏まえて、日本語習得適性テストの研究開発が本格的に始められたのは 1980 年代からである。主に名古屋大学や筑波大学で研究され、1981 年に開発された日本語習得適性テストは同年 4 月に名古屋大学総合言語センター（当時、現＝留学生センター）の日本語研修コースの学生に対して行われた。当初の目的は 6 ヶ月という限られたコースの時間の中で、クラス分けにかかる約 1 ヶ月もの期間をいかにして削減できるのかということに挑戦するものだった。初版段階では MLAT や PLAB の問題形式に基づいて作られていたが、その後信頼性、妥当性を高めるために 2 度の改訂を経て、1985 年から第 3 版が実施された。筑波大学の留学生教育センターもデータの収集に協力し、両者のデータを分析することで 1986 年頃から本格的な分析研究が始められた。

私の手元にある資料は 1993 年から 1994 年 10 月までの 3 回に渡って実施されたもののみである。1993 年に行われた適性テストでは教室での集団式テストに加えて、同じ問題をソフトウェア化し、コンピュータでも実施された。そして、この 1 回のテストでコンピュータでの実施の有効性がある程度見えたと同時に、いくつかの問題点がわかったため、問題の内容に改良が加えられた。それが 1994 年の 4 月と 10 月に行われたテスト（形式：集団式）である。しかし、残念ながら 1994 年 10 月以降、テストのソフトウェア化や更なる研究開発活動が行われたという情報や資料が公開されていないことから、研究活動は目立って行なわれていないと考えられる。

3. 日本語習得適性テストの分析結果からみる問題点

ここからは第 3 版と呼ばれている適性テストの 93 年度版から 94 年度 10 月版に移行した過程を見ながら、各問題の主旨や未だ残されている問題点などについて挙げていきたいと思う。

日本語習得適性テストは大きく分けて 3 つの項目から構成され、項目毎に数種類の問題が作られている。全部で約 90 分かかるといふ。これらの内容は大きく分けると以下の通りになる。（※問題の提出順序は以下の通りではない。）

- 1) 視覚情報処理能力測定 — a. 図形再認問題
 - b. 漢字要素抽出問題
 - c. 漢字パターン認識問題
 - d. 漢字検索問題
 - e. 音節表記問題
- 2) 聴覚情報処理能力測定 — a. 音の識別能力測定問題
 - b. 音の記憶能力測定問題
 - c. 音と形を対応させて記憶する能力を測定する問題
- 3) 文法情報解析能力測定 — a. 語構成の規則を発見する問題
 - b. 構文の規則を発見する問題

日本語習得適性を見るには、まず日本語の他の言語にはない特徴的な部分に、より早く適応できるかどうかのポイントとなる。その点で測定すべき内容がこれら 3 つに絞られたようだ。以下、上記の各項目別に見ていくが、本節では集団式で行われたものだけに留め、1993 年に行われたコンピ

ュータでのテストについては、次節で今後の可能性なども合わせて述べたいと思う。

1) 視覚情報処理能力測定テスト

日本語はひらがな、カタカナ、漢字の3種類の文字を使って表記される。(この場合ローマ字表記は除外する。)非漢字圏学習者にとってこの3つは日本語習得の始めの大きな壁となる。限られた期間で学習する場合は特に文字を覚える適性というものが鍵になる。もし、文字表記に早く適応できれば、それだけでも習得のスピードは多少なりとも速くなると考えられる。多くの日本語テキストが入門レベル以外ではローマ字表記で作られていないことからわかるように、文字を覚えなければ後々支障が出てくる事は容易に想像できる問題であるからだ。

視覚情報処理能力測定テストは全体的に94年度10月版に改訂してからのほうが、 α 係数が高くなっていて、以前と比べて妥当性は高くなったといえる。とはいっても、まだまだ課題は多く残されている。

a～eの問題をトータルで見た場合、全体的にひらがなとディクテーションテストの結果との相関は高いが、カタカナ、漢字テスト結果との相関はさほど高いとは言えない。さらに、漢字テストは課が進むにつれてテスト結果との相関が低くなっていて、入門期を越えた学習者の文字習得能力は、現行の適性テストでは測定しきれないとも考えられる。国立大学の留学生センターに所属する学生の場合、民間の日本語学校の学生に時々見られるような怠慢はあまりないと聞くので、漢字習得テストの結果が適性テストの結果と一致していなくても、それが怠けた結果からというのは考えにくい。むしろ、適性テスト自体に、インプットされる情報が多くなってきた場合に、うまく処理できるか否かを予測する性質があまりないことに原因があると考えられる。

ひらがなを覚えて、次にカタカナを覚える段階で、似て非なる文字情報が入ってきた時、どのように処理して覚えるのか、そこがカタカナ習得能力を測定する鍵となるのではないかと思う。漢字も同様で、始めに基本的でごく簡単な漢字を教わっているうちは混乱も少ないが、語数が増えてくるにつれて、似たような漢字を混同し始める。特に日本語の場合、漢字1つに対して読み方が1つにはなっていないため、余計に混乱する。そこで、新しい文字情報が次々に入ってきた場合に、それらをうまく処理できるのが漢字習得の鍵となるのではないかと思う。

文字習得の得意な学生とそうでない学生との差は何であるのか、再度検討しなおしてから、文字の適性テストの開発自体も再検討していかなければならないのではないかと感じた。

2) 聴覚情報を処理する能力の測定

日本語の音声は他の言語と比較して、複雑な音韻体系を持たず、さほど難しくないとされているが、実際は学習者の母語によって難しい音が存在する。例えば中国人、韓国人にとっての有声音、無声音の区別や、多くの学習者が特殊拍の聞き分けが出来ないことが挙げられる。これらを聞き取る能力が日本語習得において必要とされている。また、聞き取った音を文字と連携して認識することができるという能力も必要とされる。こういった能力の差が日本語習得能力の違いにつながるのと考えから、聴覚情報処理能力測定問題が作成された。しかし、93年度版まで実施していく中で問題が発生していた。

音の記憶能力測定問題や音の識別能力測定問題における問題の内容はこの3回のテストで大きく変化した。また、音と形を対応させて記憶する能力を測定する問題も94年度版で改訂された。

けれども、これらの問題はまだまだたくさんの課題を抱えている。適性テスト作成の元となるデータが少ないのも事実であるため、音の識別、音の記憶、音と図形を対応した記憶に関する能力が日本語習得のどの部分において関係してくるのかがまだ全然わかっていない。例えば、特殊拍がよく聞き取れれば会話能力と高い相関が見られるのかといった細かいデータは出ていない。まずはそこからスタートしなければ、今後新たに問題作成をしても意味がないのではないかと思う。

3) 文法構造を解析する能力の測定

日本語の文法は欧米の代表的な言語や中国語とは違った構造をもっているため、多くの学習者がその規則に慣れるのに苦労をする。例えば英語の場合、語順の基本はSVOで、主に単語自体の変化が文法関係を表す屈折語であるため、SOV言語で、且つ機能語(助詞)が付くことによって文法関係を表す膠着語である日本語に慣れる事が、英語母語話者にとって困難であることは、ある程度想像できる。しかし、本人達にとってどの程度困難であるかは個人差があるのも確かである。そういったことから、このテストは、日本語の単語の活用や、助詞も含めた語順のルールを解析する

能力があるかどうかを測るために作られた。この問題は以下のA)～I)の9つに分けられる。

- a. 語構成の規則を発見する問題
 - A) 動詞の辞書形を否定形にする問題 (2問)
 - B) い形容詞を否定形にする問題 (2問)
 - C) 動詞の辞書形をて形にする問題 (2問)
 - D) 動詞の辞書形を可能形にする問題 (2問)
- b. 構文の規則を発見する問題
 - E) 目的語のない動詞文の語順規則を見つける問題 (2問)
 - F) 目的語のある動詞文の語順規則を見つける問題 (4問)
 - G) 形容詞+名詞の規則を見つける問題 (2問)
 - H) 名詞+助詞+名詞(修飾)の規則を見つける問題 (4問)
 - I) 連体修飾の規則を見つける問題 (5問)

これらは全て、例で規則が示され、その後にある問題をその規則に当てはめて正しい答えを選択する形式になっている。英語と日本語に似せて作られた造語(ローマ字表記)とで問題は作られていて、E)～I)の問題を解くために、英語と造語の辞書(英造/造英辞書)も予め配られているので、それを見ながら答えることができるようになっている。

テスト結果のデータが94年度10月版しかないので、これを元にこのテストの信頼性などを見てみる。平均点は18.45、 α 係数は0.803と高い数値を示していて、信頼性は高いようだ。問題別に見てみると、A)、B)、C)、D)、E)、G)は正答率が高いが、F)、H)、I)はやや低い。このことから、助詞が入っている問題の中でも比較的長い文の規則を見つける問題は難度が高いことがわかった。しかし、助詞の表記が「p」=「は」「q」=「を」「x」=「の」となっていて、F)のように「p」と「q」が同じ文に出てくると見た目に紛らわしいので、構文規則を発見する能力とは別の要因から正答率を下げてしまう可能性が出てくる。こういった意図で「p」と「q」を使ったのかわからないが、できれば見た目にもっと違ったアルファベットを使って同じ問題を実施したほうがいいのではないかと思う。

以上、名古屋大を中心に開発された適性テストを見てきたが、まだまだたくさん問題や課題が残されたままになっていることがわかった。私の手元にあるデータが少ないため、何ともいえない部分も多いが、全体的に問題点が明確に示されてきた段階で研究開発が止まってしまっている感があり、非常にもったいなく思う。

この研究開発については、「本テストでは主たる目的をクラス分けのためであるのか、その後の学習のつまずきを予測するために行うのか、診断のために利用するのかなど、明確に定めないところからスタートしたところがある。基礎研究の次の段階として目的を具体的に定めたテスト作成の段階にきている。」(酒井、1995)と述べられていて、このテストの不完全性を製作者自ら指摘している。適性テスト開発の目的は前に書いたように、クラス分けを主としていたと思っていたが、他の目的は本当はないのか、具体的などころはあやふやなまま開発を進めてきたようだ。目的が具体的にしなければなるほど、もっとテストの内容も多様化しなければならなくなると考えられる。クラス分けの材料とする分には現行のテストのように、個々のテストと学習者の四技能の能力に対する相関が曖昧でも総合点からある程度妥当なクラス分けができると思われるが、もし、その後の学習のつまずきを予測するためにテストを実施するのだとしたら、もっと相関関係が明確にならないと予測は不可能である。例えば、漢字を使ったテストが漢字の習得能力にそのまま反映されていない。問題Ⅱはどの技能に対して相関が高いのか、視覚情報処理能力を予測するために作ったテストであっても、他の技能に対する予測力はないのか、他に何を予測することができるのか、等これからの課題として、個々の問題に対するもっと深い分析が求められるであろうと思う。

4. 日本語習得適性テストの実用性

4-1 就学生対象日本語学校における使用と考えられる問題点

名古屋大学で開発された日本語習得適性テストは、前にも述べたように国立大学の留学生に対して行なわれたものであり、民間の日本語学校でこのテストを行うにはもう少し違った角度で問題を作成する必要があるのではないだろうかと考え、まずは日本語学校の就学生と、大学の留学生セン

ターの留学生との違いから検討してみた。下の表は私が現在勤めている日本語学校就学生クラスと、名古屋大留学センター6ヶ月研修コースの学生についてわかる範囲で比較したものである。

	日本語学校就学生	留学生センター留学生
国 籍	中国系、韓国系で9割を占め、他はいてもクラスに数人	提携校、姉妹校からの留学生が多いため多国籍
学習環境	1クラス15～20人(25人くらいになる時もある)。基本的に全クラス同じカリキュラム。2年コースに在籍する学生が多い。教室の設備は各クラス平等ではない。	1クラス10人弱～20人。クラスによってカリキュラムが違う。6ヶ月の短期コース。
学習動機	これから日本の大学、短大や専門学校へ進学するために受験用の日本語を勉強する	すでに国の大学で自分の専門分野の研究をしていて、それを日本の大学や大学院で行うために必要な日本語を習得する
経済的事情	多くの学生がアルバイトなしに生活することが困難な状態	特に問題のない学生が多い

この比較からわかるように、就学生と大学の短期留学生とでは置かれている状況が全く違うので、名古屋大学の適性テストをそのまま使ってもその後の学習結果との相関や信頼係数が同じように出るのかという疑問点がある。その点について今後検討しなければならないが、それ以前にテスト実施自体にも大学と日本語学校との設備の差などから何らかの影響があると考えられる。もし、日本語学校で名古屋大学の適性テストを実施するとしたら、まず以下の問題をクリアする事が最低条件になると思う。

1) 問題指示文等の他国語への翻訳

大学の短期留学生は、英語が理解できることを受け入れの条件にしているため、全員英語がわかるが、日本語学校の学生の多くは高校を卒業したばかりの中国人であるため、英語が全くわからない学生もいる。韓国大学生に関しても、それほど英語力が高いとはいえない人もいるので、翻訳してあげた方が無難であると思われる。他にもロシア語などが必要となるかもしれない。

2) 外国語に精通したテスター等の準備

受験会場の各教室に受験者の母語が話せるテスターを配置するか、もしくは教室での指示に必要となることばを予め多国語でテープに吹き込んでおく必要性も考えられる。次の問題に進む時の指示や、不正行為に対する注意などする時に日本語では通じないからだ。

3) テスターの訓練と受験会場の整備

日本語学校で行う場合、学校に数十台のコンピュータが普及していれば別だが、ほとんどの学校でコンピュータは事務や教務が使う分しか置かれていない。そうすると、テストはどうしても集団式ペーパーテストで行うしかない。スライド担当、テープ担当、インストラクターの3人の連携を取るために名古屋大学のテスター達も十分な訓練をしてからテストに望み、またスライドを見る位置や教室の明るさ、スライドの明度など受験者に均一の条件下で受験してもらうために多くの工夫をしたそう。しかし、このテストを学外に持ち出して実施するとすると、テスト製作者の意図していないところで問題が発生する可能性もあり、テストの信頼性に影響を与えてしまうことも考えられる。できるだけ、問題の主旨等を理解してもらうために学校側に対して研修会を開く必要も出てくると思う。

他には、表にあるように、日本語学校の学生は漢字圏学習者が多い。中国人学習者が視覚情報処理能力測定問題を受ける必要があるのか、その点についても調べなければならない。もし、テストの結果が低かったとしても、彼らが漢字を習得する能力は非漢字圏学習者と比べて圧倒的に高いことは容易に予測できる。視覚情報処理能力測定問題の結果と四技能別能力との相関が明らかにされていない以上、漢字の問題をそのまま使うわけにはいかないのではないかなと思う。極端な話になる

が、例えば同じ地方出身の中国人2人が適性テストを受けたとする。聴覚情報処理能力測定結果と文法構造解析能力測定結果が全く同じ結果になり、視覚情報処理能力測定結果のみに差が出た場合、1人を、漢字練習をゆっくり進めるクラスに入れ、もう1人を、漢字練習を速く進めるクラスに入れるというわけにもいかないだろう。この2人の能力のどの部分に差があるということが予測できたのか、今のままでは検討もつかない。テスト結果から、クラス分けや、カリキュラムの見直しをしようとしてもこの2人に対し、どのような授業を行えば習得の助けになるかわからなければ意味がない。これは、韓国人学習者が文法構造解析能力測定テスト（問題 E）～I）を受けることにどのような意味があるのかという問題とも関わってくると思う。

以上の点からみて、現行のテストそのままでは日本語学校の就学生に対して適性テストを実施しても、結果を持って余してしまう可能性が高く、前節の最後に述べた事に関する分析の結果を出すまで、安易にテストを行ってこれをクラス分けの材料とすることは現段階では無理があると思われる。

4-2 適性テストのコンピュータソフトウェア化と実用性

93年度10月に名古屋大学、筑波大学、東北大学の3校共同でコンピュータを使った適性テストが行われた。テストの内容は93年度版の適性テストと全く同じ物である。（ただし、ペーパー版では間違いに気付いた場合に前に戻って書き直すことができた問題に関して、操作を単純化するために前に戻れないようにプログラムされた。等の多少の異同はあった。）

テストをコンピュータで実施することで、以下の5つのメリットが予測された。

- ① 受験者の反応時間など、ペーパー版の試験では測定が不可能だった項目についても測定が可能になる。
- ② 各問題の難易度についてのデータが十分に蓄積されていることが前提となるが、受験者の能力に合わせた問題を集中的に出題して、より効果的に測定が行える適応型テストにすることができる。
- ③ 従来のペーパー版では、刺激の制御にスライド、テープを用いているため、測定には最低でもスライドを制御する人間、テープを制御する人間、インストラクターの3名のテスターが必要である。さらに、様々な機関で実施される場合の受験条件を均質化するためには、あらかじめテスターを訓練しておく必要がある。コンピュータを用いて刺激の制御を行うことによって、訓練されたテスターなしでも均質な受験条件を確保することができる。
- ④ ペーパー版のテストでは、受験者の位置によってスライドの見やすさ、テープの聞きやすさが異なるなど受験条件に多少の格差が生じる。コンピュータを用いることによって、このような差を解消することができる。
- ⑤ ペーパー版の試験では、スライドとテープは全ての受験者に対して均一に提示されるため、受験者は自分のペースで試験を受けることができない。コンピュータに載せることにより、受験者は自分のペースで受験することが可能になり、この負担が軽減される。

（内山、1995）

この時の実施の目的は、コンピュータに対する慣れの違いが測定データ（得点）に影響するかどうかを明らかにすることだった。そのために、ペーパー版試験と平行して実施され、結果の比較、検討が行われた。もう1つの目的は、テスターにどのような注意が必要となるかを明らかにすることであつたため、試験実施の際に受験者の様子を観察、記録したものと、受験者にインタビューした回答とが合わせて分析された。

コンピュータはまだMicrosoft社のWindows（以下Win）が出る前だったため、海外での普及率も高かったApple社のMacintosh（以下Mac）が選ばれた。解答の入力はマウスを使った場合、マウスに慣れている受験者とそうでない受験者との差が出てしまうため、全てテンキーで入力するように作られた。1～9が解答選択用のキーで、0が確定用のキーとなった。0を押すまでは1～9のテンキーを押し直すことで解答を何度でも変えることができた。問題が提示されてから0キーを押すまでの時間が解答時間として記録された。測定時間は1人あたり64～100分だった。

受験者のテストに対する評判はペーパー版よりコンピュータ版のほうがよかった。主な理由は以下の通りであった。

① 自分のペースでできる

ペーパー版では自分が終わった後でも制限時間まで待たなければならなかったが、終わったらす

ぐに次へ進めてよかった。

② 音が聞きやすい／スライドより絵が見やすい

ペーパー版で問題点として挙げられていた、スライドとテープの問題がコンピュータソフト化によって解消された。ペーパー版では、受験者はスライドのために暗くされた部屋で画面と解答用紙の両方を交互に見なければならず、しかも着席位置によってスライドが見えにくかったりもしたが、コンピュータ版では目の前の画面から刺激が提示されるので見やすくなった。音も、ボリュームを個人ごとに設定できたので、聞きやすくなった。

③ 書く必要がないから簡単

これは、音声の同定問題に関して出た意見で、○、×という記号は日本人にとっては当たり前に書けるものであるが、受験者にとっては異同の判断の後でさらに○、×の記号の判断をしなければならなかったもので、それが負担だったということだ。

これらの意見から、コンピュータ版実施に当たり予想したメリットと同様の意見が確認された。しかし、同時に受験者からいくつか要望が上がっていた。

- ① 1時間以内で終わって欲しい。
- ② 文法が長すぎる。10問ずつぐらいに分けて、間にリスニングなどの問題を入れてもいい。
- ③ 画面の色を工夫したほうがいい。
- ④ 聞くテストの時は画面を暗くして目を使わないほうがいい。

①②に関しては、コンピュータ版固有の問題ではない。③に関しては、当時は使用端末がカラーモニターとは限らなかったために、全てモノクロの設定にしたそう。今だったら何の問題もなくカラーで作成できるだろう。④に関しては、画面上の選択肢番号を反転するという視覚的補助を加えたということだが、それが生かされなかったようだ。

テストの結果を比較してみると、ペーパー版とコンピュータ版とで全く同じというわけにはいかなかったが、ほぼ同じ結果が得られた。コンピュータへの慣れがテスト結果に影響することなかったようだ。ただし、音と形の対応記憶能力測定問題（短期）の成績はコンピュータ版のほうが悪かった。この原因としては、ペーパー版では図形に色が使われたいたものが、モノクロになっていたことと、ペーパー版ではインストラクターが音声に合わせて図形を指していたのに対し、コンピュータ版では図形の下方につけられていた点を反転させた事が結果に影響したとしている。

これらの結果と分析から、コンピュータ版のほうがペーパー版より受験者の負担が小さく、メリットが大きいと考えられた。

適性テスト自体は、コンピュータ版の結果を受けて94年に改訂したものが2度実施されたが、いずれもコンピュータ版にはなったという報告がない。94年度版の実施で一番問題となった漢字検索問題に関していえば、今これをコンピュータ版にすることは簡単な作業である。作問者は当時この問題を作った時に、漢字を1つずつ画面に出していけば飛ばすことなく順番通りに答えさせられると考えていたようだが、一気に全部画面上に出して、答えるまで次に進めないようにプログラムすれば、このまま同質の問題としてデータを出すことも可能である。

約10年前に、モノクロ画面のMacでこれだけのことができたことを考えれば、今WinXPを使ってもっと問題を多様化させることもできるであろう。コンピュータ自体のデータ処理速度も格段に速くなっていて、メモリーも大きいものが標準で組み込まれているので動作も安定しているので、データが重くてもエラーやフリーズは起こりにくい。極端な話、必要であれば刺激に動画を使った問題も作ることができる。

また、テスト結果から得られる得点や反応速度のデータも、そのままコンピュータが勝手に計算して出してくれるようにプログラムしておけば、受験者数が多くてもデータ分析にかかる時間が削減できる。

さらに、ニーズがあればオンラインでこのテストを実施することも可能である。海外のビジネスマンやこれから日本語を学ぼうとする学生達が、オンラインでこのテストを受けた後、結果をこちらに送信すれば、分析結果を追ってメール等で通知できるようにする。また、日本語教師の立場から見てもオンライン適性テストの利用価値は高いと考えられる。例えば、海外の大学の日本語科で

どのような授業を行うか考える材料として、学生達にこのテストを受けてもらうこともできる。

ただし、コンピュータ版適性テストは日本語学校で行うのはまだまだ難しいだろう。日本語学校で学生用のコンピュータルームを持っている学校はほとんどないといってもいいし、これからもそれほど増えないと思う。ペーパー版で行うことになった場合には、せめてスライドではなく PowerPoint を使って視覚の刺激を与える際にテスト結果にマイナスの影響を与えないように工夫したい。

5. 今後の課題とまとめ

今回、名古屋大学のデータをもとに色々見てきた結果、適性テスト開発の難しさがわかると同時に問題点がはっきりした分、これからさらに研究を続けていく意欲も湧いてきた。名古屋大のテスト自体のこれからの課題については第3節で述べた通りだが、私自身に課す課題も明確になった。

元々日本語教育研究における私の専門分野は音声学だったため、今回名古屋大学適性テストの聴覚情報処理能力測定問題の分析結果から考えさせられることがたくさんあった。聴覚情報処理能力測定問題のところでも触れたが、音声識別問題に関して、音の聞き取りが得意か否かで日本語習得においてどんな能力が測れるのか、この結果からはわからなかったが、これは非常に興味深い問題だと思った。よく、「あの学生は耳（聴解能力）が弱いから、いつまで経っても覚ええないのよ」などという言葉聞くことがあるが、果たして、単音を聞き分ける能力と、日本語習得能力はどう関係しているのか。また、「ツ」と「チュ」の聞き分けができなくても、清濁の聞き分けができる場合に、それが習得にどのような影響を与えているのか、何らかの妨げになっているのかいないのか。色々調べてみる必要が出てきた。この問題がクリアになれば聴覚情報処理能力測定問題を作ることは難しいだろう。

それから、もう1つ調べてみたいことは、口真似が上手な人は語学習得能力も高いのかということである。数十年前に大阪 YMCA が英語の適性テストを開発した時、口真似テストも入れていたそう。そのテストの信頼係数のデータがないのでよくわからないが、もしこれも日本語習得の適性を測るのに有効な手段であれば研究してみたい。

これからも引き続き、適性テストの開発研究を続けていきたいが、テスト全体を開発するのは1人ではとても無理である。テーマを決めた時に何とか少しでも問題を作ってみてデータを出そうと意気込んでいたのだが、この開発には少なくとも5名の専門家の協力体制が必要であると思った。文法を専門とする人、文字習得を専門とする人、音声学を専門とする人、心理学を専門とする人、評価法を専門とする人である。更に研究が進めば、もう一名コンピュータプログラミングを専門とする人も必要となるかもしれない。私の専門は音声学であるが、上記のとおり、音声の適性問題作成には厚い壁が立ちはだかっているため、全く完成の目処は立たないが、あきらめずに研究していきたいと考えている。

参考文献

- (1) 日本語教育学会 (1991) 『日本語テストハンドブック』大修館書店
- (2) 石田敏子 (1992) 『入門日本語テスト法』大修館書店
- (3) 大坪一夫、酒井たか子、藤原雅憲、村上京子、行広隆次、他 (1995) 『パーソナルコンピュータを用いた外国人日本語学習者のための適性テストの開発』1993、94 年度文部省科学研究費助成金総合研究 (A) 研究成果報告書